The Astrophysical Journal, 714:476–486, 2010 May 1 doi:10.1088/0004-637X/714/1/476


# A NEW APPROACH TO ASTROPHYSICAL SPECTRA: THE COMPLETE EXPERIMENTAL SPECTRUM OF ETHYL CYANIDE ($CH_3CH_2CN$) BETWEEN 570 AND 645 GHZ

Sarah M. Fortman, Ivan R. Medvedev, Christopher F. Neese, and Frank C. De Lucia
Department of Physics, Ohio State University, 191 W. Woodruff Ave., Columbus, OH 43210, USA; fcd@mps.ohio-state.edu
*Received 2009 December 15; accepted 2010 March 15; published 2010 April 12*

## ABSTRACT

There is a general consensus that many of the unidentified features in astrophysical spectra are due to low lying excited vibrational and torsional states of a few molecules—commonly referred to as the astrophysical weeds. This is a challenging spectroscopic problem not only because there are many such states, but also because these states are often highly perturbed and difficult to analyze. We have previously described an alternative approach based on experimental, intensity-calibrated spectra taken at many temperatures. In this paper, we describe the procedures and results obtained with this approach for ethyl cyanide, strategies for archiving and disseminating these results, and the prospects for using these results to reduce the confusion limit in the powerful new observatories that are coming online.

*Key words:* astronomical databases: miscellaneous – ISM: molecules – molecular data – submillimeter: ISM – techniques: spectroscopic

*Online-only material:* color figures, machine-readable tables, tar.gz file

## 1. INTRODUCTION

With the advent of powerful new submillimeter (SMM) astronomical instruments such as Herschel, ALMA, and SOFIA, many astrophysical problems are making a rapid transition from being telescope limited to being limited by spectroscopic knowledge. Indeed, in hot molecular cores, even before the advent of the latest generation of telescopes, as many as 50% of the observable lines were unidentified. There is a strong consensus that the large majority of these unidentified lines arise from the low lying excited vibrational and torsional states of a few molecules, the astrophysical weeds.

The spectroscopic problem is not only that there is a large number of such states, but more importantly for many of these states it is also difficult to assign and build quantum-mechanical (QM) models because of the strong perturbations and interactions among them. We have proposed an alternative approach (Medvedev & De Lucia 2007) to complement the traditional bootstrap assignment and modeling. This approach is based on complete, intensity-calibrated spectra, taken over a range of temperatures.

In this paper, we describe the results of a study of ethyl cyanide in the frequency range 570–645 GHz. This molecule was chosen because it is one of the astrophysical "weeds," especially in hot cores where the temperature gives rise to detectable excited state populations (Mehringer et al. 2004) and because it is a well behaved, relatively rigid molecule for which we can use theoretically calculated intensities as benchmarks for our new approach. With the establishment of calibration and verification techniques described in this paper, we are working to extend this approach to the other astrophysical weeds, including less rigid molecules such as methyl formate and methanol for which calculated intensities may be problematic.

This work provides the astrophysically required transition frequency, transition strength, and lower state energy without line assignment or the use of a QM model. This paper also discusses an alternative approach that remains useful in the limit of crowded and blended spectra. The specific experimental and theoretical techniques used are described, as are strategies for archiving these results so as to optimize their usefulness to the astrophysical community.

Figure 1 shows a comparison between a complete laboratory intensity-calibrated spectrum and the results of the usual QM model-based catalog. (Pickett et al. 1998; Müller et al. 2005). Because of the amount of information contained in this spectrum, this comparison can only be shown after a series of spectral expansions, the last showing ∼0.25% of the spectral range. The significant point here is not only that many of the spectral lines are missing from the catalog, but also that many lines that are significantly stronger than many of the catalog lines are included in the experimental spectrum.

Figure 2 shows a quantitative overview of this comparison. Of the 9962 strongest lines in the experimental spectrum, only 695 are in the catalog, while 386 of the lines that are in the catalog are not among the 9962 strongest lines. While the strongest lines in the spectrum belong to the ground state, many lines that are significantly stronger than catalog lines do not.

## 2. BASIC EQUATIONS

We have previously shown that with intensity-calibrated data taken over a range of temperatures, it is possible to obtain the astrophysically significant lower state energy and transition strength without the need for QM assignment (Medvedev & De Lucia 2007). Briefly, the integrated absorption coefficient for a rotational transition is given by (Gordy & Cook 1984)

$$\alpha_{l\to u}(T) = \nu n(1 - e^{-h\nu/kT})\frac{8\pi^3}{3ch}\sum_{i=x,y,z}\left|\mu_{i,l\to u}\right|^2\frac{g_l e^{-E_l/kT}}{Q}, \quad (1)$$

where $\nu$ is the line frequency, $T$ is the temperature, $\sum_{i=x,y,z}|\mu_{i,l\to u}|^2$ is the square of the dipole matrix element, $g_l$ is the degeneracy of the lower level, $E_l$ is the lower state energy, and $Q = \sum_{n=0}^{\infty} g_n e^{-E_n/kT}$ is the partition function.

If an unassigned line (Line 1) and an assigned line (Line 2) are measured and their peak absorption coefficients, normalized for Doppler broadening from Equation (1), are ratioed, then

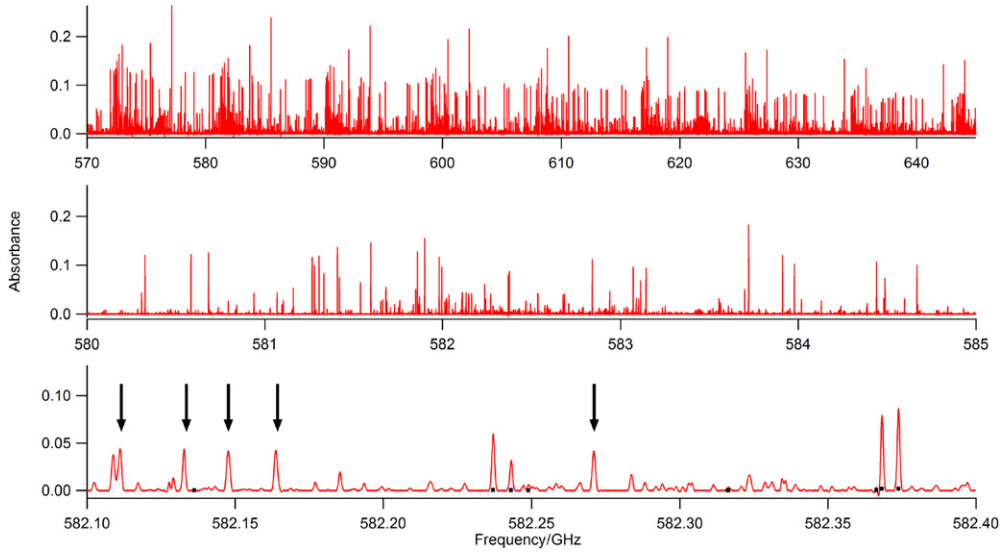


**Figure 1.** Spectrum of ethyl cyanide at 300 K, shown as a series of spectral expansions. In the lower panel (expanded 2× vertically) a comparison is made with a catalog based on a QM model (the dots indicate frequency and intensity of the catalog lines). The arrows show lines that are not in the catalog that are significantly larger than other lines that are in the catalog.

(A color version of this figure is available in the online journal.)

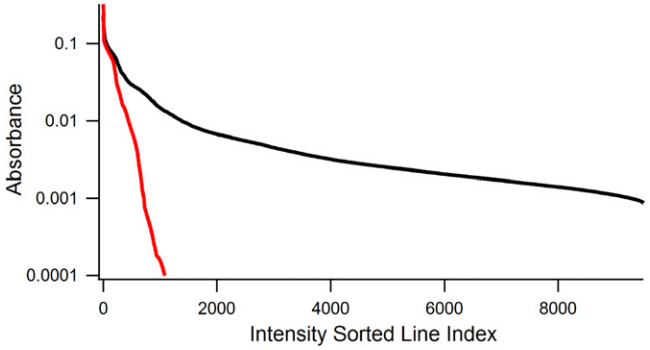


**Figure 2.** Comparison of the intensities of the transitions (sorted according to absorbance) for the experimental lines of Figure 1(upper black points) and transitions from a QM catalog (lower red points). The densities of points are so high as to appear as continuous lines. The two traces separate at an intensity that is down by the population factor for the first two excited vibrational states ($v_{13}$ at 207 cm$^{-1}$ and $v_{21}$ at 213 cm$^{-1}$) at 300 K.

(A color version of this figure is available in the online journal.)

$$[\alpha_{1,\mathrm{peak}}(T)/\alpha_{2,\mathrm{peak}}(T)][(1-e^{-h\nu_2/kT})/(1-e^{-h\nu_1/kT})] = Ce^{-(E_1-E_2)/kT}, \quad (2)$$

with

$$C = \frac{\left[\sum_{i=x,y,z} \left|\mu_{i,l\to u}\right|^2 g_l\right]_1}{\left[\sum_{i=x,y,z} \left|\mu_{i,l\to u}\right|^2 g_l\right]_2}. \quad (3)$$

If the natural log of the left side of Equation (2) is plotted against $1/T$, Figure 3 results, which can be analyzed for the only unknowns, the lower state energy of the unassigned line, $E_1$, and

$$\left[\sum_{i=x,y,z} \left|\mu_{i,l\to u}\right|^2 g_l\right]_1. \quad (4)$$

This temperature-independent term is the same as ${}^{x}S_{JJ'}\,\mu^2$ (Townes & Schawlow 1955), $\mu_g^2\lambda_g(J,\tau,J',\tau')$ (Gordy & Cook 1984), and $S_{ij}\mu^2$ in Splatalogue[1]. For consistency with Splatalogue and compactness of notation, we will adopt the notation $S_{ij}\mu^2$ for the quantity in Equation (4). Although not explicit in Townes and Schawlow or Splatalogue, $\mu$ refers to the component of the dipole moment along the principle molecular axes that is responsible for the transition, and $S_{ij}$ is a unitless quantity (the transition strength factor), which is readily calculated from a QM model. These notations also use various standard notations to indicate the transitions involved. In order to separate $S_{ij}$ from $\mu^2$, the transition involved must be assigned. Thus, we treat $S_{ij}\mu^2$ as a single quantity that we refer to as the transition strength.

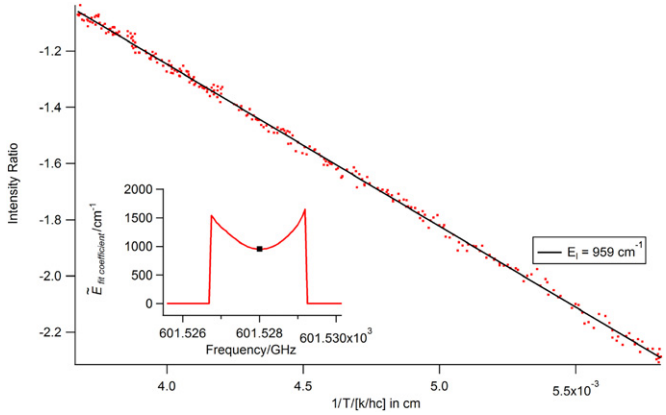


**Figure 3.** Plot in $1/T$ space to obtain the lower state energy and line strength directly from the measured peak absorbance for the unassigned line at 601528 MHz. The inset shows the results for the calculated energy for the frequency point-by-point calculation discussed in Section 5.2. The dot on the inset corresponds to the true energy (959 cm$^{-1}$) of the transition.

(A color version of this figure is available in the online journal.)

## 3. EXPERIMENTAL DETAILS

The absorption cell used in our measurements was a 6 m long, 15 cm diameter, thermally insulated aluminum pipe with

[1] http://www.splatalogue.net

a fiberglass insulated wooden enclosure. Ten independently controlled (Omega CN7500) heaters (Thermocraft), with thermocouple sensors, were used to continuously ramp the temperature. Cooling was achieved convectively by means of liquid nitrogen filled finned tubing inside the thermal enclosure along the entire length of the cell. Initially, the cell was cooled to −40 °C at which point heaters were turned on to maintain uniform temperature distribution as the temperature was ramped to 140 °C over a period of ∼3 hr.

The uniformity of the temperature distribution was monitored and recorded continuously during the measurements. The largest temperature differential was 6 K in the regions close to the absorption cell windows, which had manually controlled heating tapes wrapped around them to minimize this effect.

The spectra were recorded using the Fast Scan Submillimeter Spectroscopic Technique (FASSST; Medvedev et al. 2004; Petkie et al. 1997). In the implementation used here a times 48 solid state multiplier chain (Virginia Diodes) was driven by a rapidly sweeping (Agilent E8256D) synthesizer. This approach has several differences from the traditional backward wave oscillator-based FASSST. Improved linearity of the frequency sweep and the thermal stability of the power output are the two most important because they are critical for faithful line shape reproduction and intensity calibration. We carefully checked the power output for the presence of undesired harmonics and were pleased to discover that the strongest "parasitic" multiplication product was down by a factor of 2000 and was below our detection limit for the short integration time and low sample pressure used in our experiment. The radiation was detected by a liquid helium cooled InSb bolometer (QMC).

Overall, 271 spectral scans were recorded. Each spectrum covered the 570–645 GHz region with a data interval of 62.5 kHz, chosen to provide about 10 points per Doppler-limited line width. To minimize the temperature and sample pressure variation over a single scan, the integration time per frequency resolution element was chosen to be 30 $\mu$s. Thus, each scan was recorded in ∼38 s. The fast scan speed was required because Equation (2) assumes that the temperature, sample concentration, and pressure remain constant during the 75 GHz sweep. While longer integration times would have produced a larger signal-to-noise ratio (S/N), the sensitivity of the system was more than adequate, and the methods for the recovery of the spectroscopic data described below effectively integrate over all the scans.

We recorded spectra with sample pressure of about 0.5 mtorr. At this pressure, the experimental line width was within 5% of the calculated Doppler line width. To maintain sample stability over the duration of the experiment, we conducted our measurements in a flow regime. To preserve the natural line shape, we recorded spectra with significant post-detection bandwidth. The high pass limit was set by the AC coupling of the preamplifier, while the low pass was set to match the Nyquist frequency of the digitizer. To recover the absolute magnitude of the SMM power, lost due to the AC coupling, we amplitude modulated the radiation by means of an optical chopper. About every 20 scans, a baseline scan was recorded with the chopper on, so that we were able to faithfully measure the baseline variations with temperature.

## 4. DATA PROCESSING

The first step in processing the spectral information was the conversion of each spectrum to naperian absorbance. We did this by removing the sharp lines and numerically fitting the remaining broader baseline features in each spectrum to a polynomial spline function. We then subtracted the resulting splines from each spectrum and divided by the appropriate total power function obtained from the chopper scans to calculate the transmission function. By taking the natural logarithm of the transmission, we obtained the naperian absorbance that is proportional to the optical path length and the absorption coefficient. Since we ran at low pressures, there was little pressure broadening, and Gaussian fits gave us very good intensity measurements. For lines that are close in frequency (less than two line widths apart), we fit for all of the lines simultaneously using multiple Gaussian peaks.

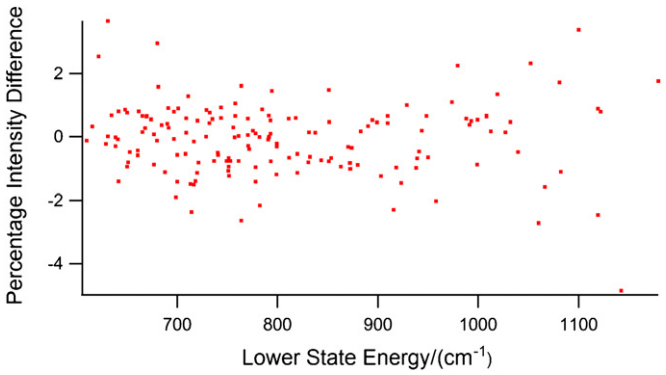


**Figure 4.** Fractional residuals of the 163 reference lines used to obtain the spectroscopic temperature. The rms of these fractional residuals is 1%.
(A color version of this figure is available in the online journal.)

Because our results depend upon the quality of the reference lines in the ratio, we chose a subset of 163 assigned lines (Pearson et al. 1994; Pickett et al. 1998), based on their strength and isolation from other lines. These reference lines include lines whose absorbances vary between 0.15 and 0.019. While, in principle, we only require one reference line, this redundancy allows us to check and test assumptions that underlie this new approach. These assumptions include that chemistry and temperature are constant over a sweep, that the molecules in the cell can be characterized by a temperature, and that there are no frequency-dependent effects (e.g., standing waves). Additionally, because these lines have known lower state energies and transition strengths, they (as well as the other assigned lines) can be used to test the accuracy of the approach, subject to the caveat that they may be overlapped by unknown and unassigned lines.

The thermocouples give us a good approximation of the temperature, but end effects and other variations still exist. Since the reference lines are assigned and have Doppler width $\delta\nu_D$, we can use a fit of their measured peak absorbance,

$$A_{\rm peak} = L\alpha_{\rm peak}(T) = \frac{nL}{Q}\frac{8\pi^3}{3ch}(1-e^{-h\nu/kT}) \times S_{ij}\mu^2 e^{-E_l/kT}\sqrt{\frac{\ln(2)}{\pi}}\frac{\nu_0}{\delta\nu_D}, \quad (5)$$

to obtain the spectroscopic temperature $T$ and $nL/Q$. By fitting for this latter parameter, we make no assumptions about the stability of the length of the cell or the number density of the sample between runs—only that they are constant over the relatively short period (38 s) of a single scan. Figure 4 shows the residuals of this fit with an rms deviation of 1%. These are sorted by lower state energy to show that the assumption of a single temperature for our cell is valid. Similar sorts by

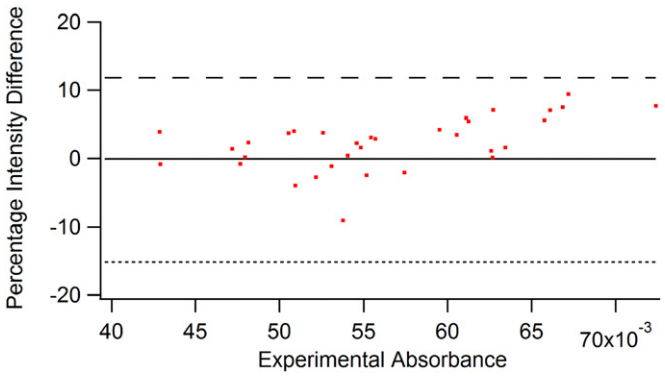


**Figure 5.** Intensity fit for the stronger b-type transitions, with the adoption of the b-type dipole moment to be 1.31 D. The dashed line at the top would be the location of the zero for this graph if the value of (Heise et al. 1974) were used and the lower dashed line would be the location of the zero if the value of (Laurie 1959) were used.

(A color version of this figure is available in the online journal.)

**Table 1**
Intensities and the Ratio of the Values of the a-Type and b-Type Dipole Moment

| Reference | $\mu_a$/D | $\mu_b$/D | $\mu_a/\mu_b$ | $(\mu_a/\mu_b)^2$ |
|---|---|---|---|---|
| Laurie 1959 | 3.78 | 1.38 | 2.74 | 7.52 |
| Heise et al. 1974 | 3.85 | 1.23 | 3.13 | 9.80 |
| This Work | 3.85 | 1.31 | 2.94 | 8.64 |

intensity (to check for system linearity) and frequency (to check our assumption about the effective cell length being constant) also produced null systematic results. The 1% accuracy of this fit holds each of these tests to a rather high standard.

One of the surprising results of this work is that we found that the values of the a- and b-dipole moment components derived from Stark measurements (Heise et al. 1974; Laurie 1959) and used in the catalogs were not consistent with our results. While we expected that there might be intensity issues for less rigid molecules such as methanol or methyl formate, it is usually assumed that intensities based on Stark measurements are very good for more rigid molecules. Inspection of Table 1 shows that whereas the a-type dipole moments are in relatively good agreement (2%), the b-types vary such that the ratio of the two components differs by 14%, and the square (to which our intensity measurements are sensitive and which are of significance to astrophysical calculations) differs by 30%.

Accordingly, we adopted the a-type dipole moment of Heise et al. (1974), which makes our a-type results consistent with the current catalogs, and adjusted our b-type value so that we obtained the correct intensity ratio between a-type and b-type transitions. The results of this are shown in Figure 5. The dashed line at the top would be the location of the zero for this graph if the value of (Heise et al. 1974) were used and the lower dashed line would be the location of the zero if the value of (Laurie 1959) were used.

In addition to the spectroscopic temperature and calibration of the dipole moments used in astrophysical calculations, the fit of the predicted intensities to the measured absorbance spectra returns the scaling parameter that is equal to the product of the number density of molecules in the absorption cell $n$ and the length of the cell $L$, divided by the rovibrational partition function $Q$. Figure 6 shows the graph of this parameter against spectroscopic temperature. As temperature gets higher, $nL/Q$ is

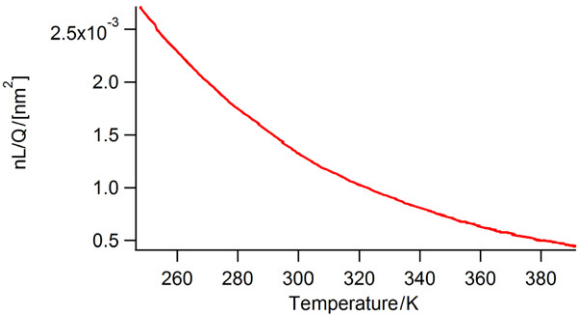


**Figure 6** $nL/Q$ parameter plotted against spectroscopic temperature. The parameter is obtained by the least-squares fitting of the predicted intensities of the 163 ground state transitions of ethyl cyanide to the experimental absorbance spectra.

(A color version of this figure is available in the online journal.)

getting smaller due to the decrease in number density as well as the increase in the rovibrational partition function value.

Scaling of every experimental absorbance spectrum by the corresponding $nL/Q$ parameter and fitting of the function in Equation (2) to the resulting scaled peak intensities is equivalent to the fitting depicted in Figure 3 followed by the averaging over all reference lines described in what follows (Section 5.1). The former method is computationally simpler, but in this project we resorted to the latter approach to have additional flexibility and to investigate the accuracy of our individual experimental data.

Division of the absorbance spectra by the $nL/Q$ parameter effectively converts them to absorption cross-section spectra, with an additional scaling by the rovibrational partition function. Since the exact value of the rovibrational partition function is scenario dependent and is not always known, scaling by it provides for a convenient method of cataloging experimental data described below in Section 5.2.

## 5. ANALYSIS APPROACHES

The goal of this work is to provide a means for the calculation of astrophysical spectra over the range of astrophysically important temperatures. We will see below that there are two distinctly different approaches. These approaches provide consistent results, but vary in convenience of use in the prediction of spectra and their uncertainty in regions with blended lines. They are as follows.

1. Logarithmic fits to $1/T$ for individual lines in the context of Equation (2) and Figure 3 to obtain the usual catalog values for the transition strength and lower state energy.
2. A logarithmic fit to $1/T$ for each frequency data point and the use of these fits to predict the spectra (including the structure of blends) via the interpolation and extrapolation of this fit.

We will show below that at the line centers of unblended lines both approaches yield virtually identical results.

### *5.1. 1/T Fits to Individual Lines*

Because of issues associated with baseline and S/N, we chose to analyze the 3355 strongest lines by ratioing each of them with the 163 reference lines, about 0.5 million analyses as represented by Equation (2) and Figure 3. This selection of 3355 lines for analysis was based on an absorbance threshold of 0.0025. While the spectrum of Figure 1 contains many more lines, as we will

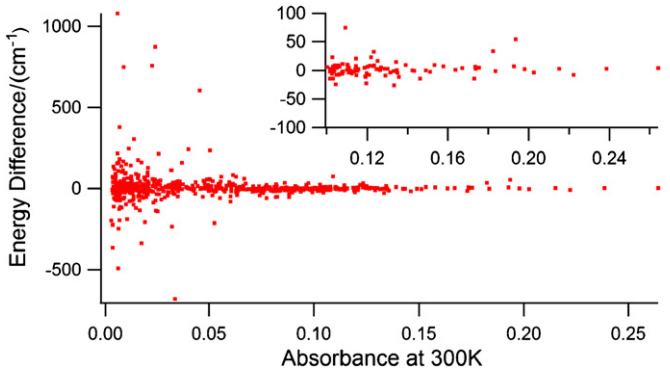


**Figure 7.** Comparison for the 585 assigned lines whose strengths place them among the 3355 strongest lines in this spectral region between their experimentally derived lower state energies and those calculated from their QM assignment. The inset expands the energy difference axis for the stronger lines.

(A color version of this figure is available in the online journal.)

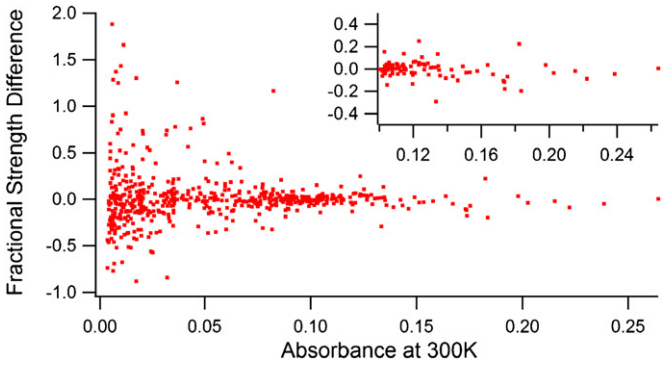


**Figure 8.** Comparison for the 585 assigned lines whose strengths place them among the 3355 strongest lines in this spectral region between their experimentally derived line strengths and those calculated from their QM assignment. The inset expands the strength difference axis for the stronger lines.

(A color version of this figure is available in the online journal.)

see below for weaker lines issues of blends begin to become important, as does the strength of the lines not only in relation to noise, but also in relation to baseline fluctuations.

We chose the pressure in this experiment to be very low (∼0.5 mtorr) primarily to avoid having lines with large absorbances that would limit our ability to accurately obtain absorption coefficients for the strongest lines. Since we can get much improved results on the lines below this 0.0025 cutoff by redoing the experiment at a higher pressure and because 3355 lines represent a very large increase in the spectral data base, we chose this particular cutoff in intensity.

A comparison of the fits, to each of the 163 reference lines, for each of the 3355 lines was made to identify reference lines that might be contaminated by overlaps with unassigned lines or other problems. The 163 reference lines that we have chosen at the end of this process appear to be contamination free, and a wide variety of reference choices from within this set give virtually identical results. We have retained this large reference set because the overall processing time was not excessive and because it provides a better opportunity to evaluate the results.

#### *5.1.1. Accuracy of Lower State Energies*

The most direct means of determining the accuracy of the lower state energies is to compare energies of assigned lines calculated from the experimental data with QM energies. The simplest approach is to compare the 585 assigned lines that are among the 3355 strongest lines, using an average across the 163 possible reference lines for each comparison (Figure 7 results). There are a number of obvious outerliers, which can be traced to cases for which unassigned lines, often much stronger, overlap lines in the QM catalog. Inspection of the 93 strongest lines (displayed in the inset) shows an rms scatter of only 14 $cm^{-1}$. Much of the contribution to this rms deviation comes from a few outerliers, which appear to be assigned lines with overlaps. The large outerliers in the remainder of the graph are almost certainly due to overlaps with unassigned lines. An important conclusion from this graph is that calculated catalogs are not only incomplete, but also contain lines that are overlapped by stronger, unassigned lines.

#### *5.1.2. Accuracy of Transition Moments*

A similar comparison, shown in Figure 8, can be made for transition moment, but here a more appropriate measure is fractional uncertainty. Again we see a strong correlation between the agreement and line strength. Problems associated with accidental overlaps with lines that are not in the catalog and experimental error for weaker lines both contribute to this correlation. While the higher fractional uncertainty of the lines with smaller transition moments is expected, it is also true that from an astronomical point of view a larger fractional uncertainty in a weak line is less significant.

#### *5.1.3. Accuracy of Calculated Spectra*

Because of the contamination of the comparisons shown in Figures 7 and 8 from overlaps from lines not included in the catalogs and because of the correlations among the parameters, we need an alternative approach to the estimation of the error produced by the experimental method. We start with the observation that the uncertainties in the observed spectra themselves are closely related to the answer to this question. Within the range of temperatures observed, because the fit line in Figure 3 averages over a large number of different temperatures, this fit line should produce a spectral intensity that is somewhat better than the measurement at a particular temperature. In Figure 4, we showed that for a selection of unblended lines, whose strength ranged from an absorbance of 0.15 to 0.019, the rms error in their fractional uncertainty was 1%. Consider two limits as follows.

1. As lines become weaker, the impact of noise and baseline becomes proportionally stronger. If these set the limit, then the absolute error is dependent on the strength of the line. In this case, because the weakest of the 3355 lines are about 1% of the strongest, this would argue that they have an error of 25%. While this is a very pessimistic view, since these same lines would be correspondingly weak in an astrophysical spectrum, the need for accuracy in the model would also be reduced by a factor of 25.
2. However, if the error were set entirely by the calibration of the absolute power in the system, then the error in the calculated spectral intensity would be independent of spectral strength and all spectral lines would have an uncertainty of 1%.

While spectral overlap and the relatively few weak lines that are both in the catalog and among the 3355 lines preclude a direct measure (similar to that shown in Figure 4 for strong lines), it is our estimate that the accuracy of the calculated spectral intensity

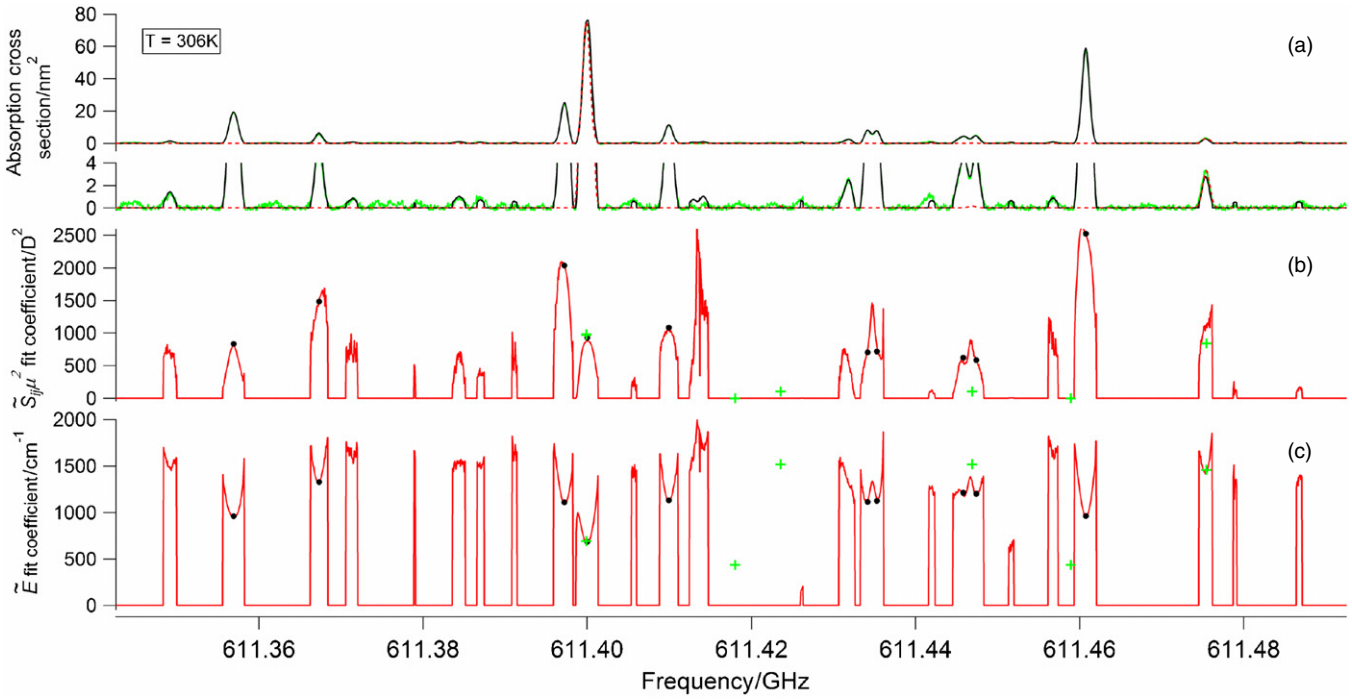


**Figure 9.** $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients calculated by fitting Equation (9) to experimental data. Panel (a) shows the spectrum calculated using Equation (9) and corresponding $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients (black trace), together with experimental data (green trace) and the catalog predictions for the ground state transitions (red dashed trace). Black dots in panels (b) and (c) are, respectively, the transition strengths and lower state energies calculated using $1/T$ fits to individual lines. Green crosses in panels (b) and (c) are, respectively, the transition strengths and lower state energies from the QM catalog.

(A color version of this figure is available in the online journal.)

will range from ∼1% for the stronger lines as shown in Figure 4 to perhaps 10% for the weakest of the 3355 lines.

### 5.2. Global Fits to the Entire Spectrum in Frequency Space

For blended lines the analysis for line parameters such as lower state energy and transition strength can be challenging, and often without meaning. However, the experimental data contain astrophysical information that can be extracted without the assumption that the spectra are made up of a collection of lines.

To realize this approach, we took advantage of the fact that spectral lines in our data are very near to the Doppler limit. If we introduce the Doppler lineshape function

$$g(\nu) = \sqrt{\frac{\ln(2)}{\pi}}\frac{1}{\delta\nu_D}e^{-\ln(2)\left(\frac{\nu-\nu_0}{\delta\nu_D}\right)^2} \tag{6}$$

to the peak absorbance relation of Equation (5), the absorbance as a function of frequency becomes

$$A(\nu) = \frac{8\pi^3}{3ch}\sqrt{\frac{\ln(2)}{\pi}}\frac{nL}{Q}\frac{\nu_0\left(1-e^{-\frac{h\nu_0}{kT}}\right)}{\delta\nu_D} \times S_{ij}\mu^2 e^{-\frac{E_l}{kT}}e^{-\ln(2)\left(\frac{\nu-\nu_0}{\delta\nu_D}\right)^2}, \tag{7}$$

with the Doppler width (HWHM) given by

$$\delta\nu_D = \sqrt{\frac{2N_a k\ln(2)}{Mc^2}}\sqrt{T}\nu_0 = W\sqrt{T}\nu_0, \tag{8}$$

where $N_a$ is Avogodro's number, and $M$ is the molecular mass.

The absorbance normalized by the $nL/Q$ factor becomes

$$\frac{A(\nu)}{nL/Q} = \frac{8\pi^3}{3ch}\sqrt{\frac{\ln(2)}{\pi}}\frac{1}{W}\frac{\left(1-e^{-\frac{h\nu_0}{kT}}\right)}{\sqrt{T}}S_{ij}\mu^2 e^{-\frac{E_l}{kT}}e^{-\frac{\ln(2)}{W^2T}\left(1-\frac{\nu}{\nu_0}\right)^2} = K\frac{\left(1-e^{-\frac{h\nu_0}{kT}}\right)}{\sqrt{T}}\tilde{S}_{ij}\mu^2 e^{-\frac{\tilde{E}(\nu)}{kT}}, \tag{9}$$

with

$$\tilde{E}(\nu) = E_L + k\frac{\ln(2)}{W^2}\left(1-\frac{\nu}{\nu_0}\right)^2. \tag{10}$$

The spectral data can be represented as a two-dimensional array of frequency- and temperature-dependent spectral intensities. According to Equation (9) every frequency slice of the data can be fitted with two parameters $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$. On the line center, when $\nu = \nu_0$, $\tilde{E}$ equals the lower state energy, while $\tilde{S}_{ij}\mu^2$ corresponds to the line strength. Off of the line center, the meanings of $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ are less physical, but Equation (9) is still a valid fitting function for describing the spectral intensity. The parameter $\tilde{S}_{ij}\mu^2$ only differs from $S_{ij}\mu^2$ because of its correlation with $\tilde{E}$ and the complications of blends.

After scaling the absorbance spectra by the corresponding $nL/Q$ parameter (see Section 4), we fit Equation (9) to the resulting absorption cross-section data and obtain $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients for each of the 1.2 million frequency points. Figure 9 shows graphs of $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients as a function of frequency for a portion of our spectral range. The top panel of Figure 9 shows spectrum, calculated using Equation (9) and corresponding $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients (black trace), together with experimental data (green trace) and the catalog predictions

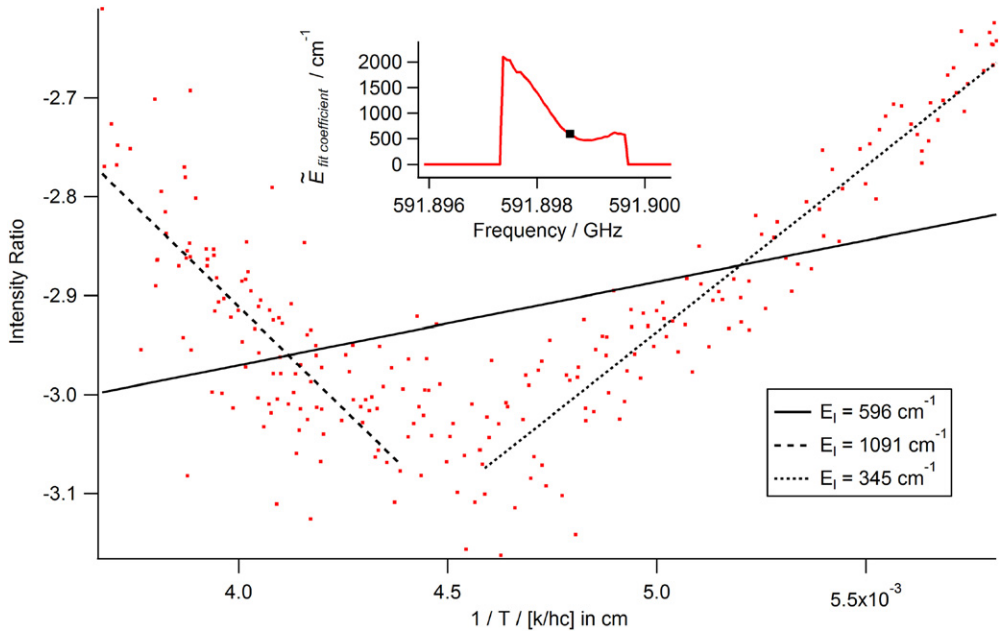


**Figure 10.** Impact of a blended line on both the $1/T$ fits and the frequency point by point fit. The $1/T$ fit is particularly interesting in that it shows the transition from the dominance of the higher energy line at high temperature to the dominance of the lower energy line at low temperature.
(A color version of this figure is available in the online journal.)

for the ground state transitions (red dashed trace). Differences between experimental (green) and calculated (black) spectra can be noticed only at 20× vertical magnification, which corresponds to ∼0.5% error in intensities of the strongest spectral lines. Despite the fact that many transitions are blended, fitting of Equation (9) derived for a single spectral line to experimental spectra allows us to reproduce data with an exceptional accuracy.

Black dots on the bottom two panels of Figure 9 are the lower state energies and transition strengths calculated using $1/T$ fits to individual lines (Section 5.1). On the line center, it should not be too surprising that two approaches produce nearly identical results. Green crosses correspond to the predicted lower state energies and transition strengths of the ground state transitions found in the online spectroscopic databases. The green crosses that do not correspond to lines in the experimental data result from the inclusion in the spectral catalogs of very weak ground state lines. Conversely, a number of stronger experimental lines correspond to lines that are not included in the ground state catalogs. Many of the transitions rejected by the $1/T$ fits are faithfully reproduced by the global fitting approach, thereby providing useful information to the astrophysical community for blended lines.

The shapes of the fit parameters as a function of frequency, especially for the "energy" $\tilde{E}$, are particularly interesting and useful. For single unblended lines, the shape away from the line center is symmetrical and reflects the additional contribution from the exponential in the Gaussian lineshape as defined in Equation (10). The inset in Figure 3 is an example of this. However, in the case of blends, both the $1/T$ plot ratios as in Equation (2) and the shape of $\tilde{E}$ (and $\tilde{S}_{ij}\mu^2$) show the impact of the overlapping lines. Figure 10 shows both of these effects for the unassigned blend near 591899 MHz.

Since most astrophysical applications call for spectra at lower temperatures, we decided to include 77 K predicted spectra into our fitting algorithm. First, we simulated the 77 K spectrum of ethyl cyanide based on the spectroscopic data available in the spectroscopic databases (Müller et al. 2005; Pickett et al. 1998; shown as dashed red trace in the top panel of Figure 11). We then added ∼1 MHz surrounding each ground state spectral line at 77 K to the experimental data and fitted Equation (9) to the combined data set. Because there is only one 77 K data set, we weighted it 10 times higher than the 271 individual experimental spectra in the analysis.

The strongest lines in the room temperature spectrum of ethyl cyanide are a-type ground state transitions with lower state energies ∼500–600 cm$^{-1}$, while the strongest lines in the 77 K spectrum are the ground state b-types with the lower state energies ∼100 cm$^{-1}$. Due to the temperature dependence of the rovibrational partition function, the Doppler line widths, and the Boltzmann factors, the 77 K strongest b-types will have ∼10 times higher spectral intensities than the strongest room temperature a-types for the same molecular number densities in the absorption cell. Depending on experimental conditions, lines from the lowest torsional state (∼200 cm$^{-1}$) of ethyl cyanide could have an appreciable S/N at 77 K. Thus, leaving frequency space between ground state lines in 77 K spectrum out of the fit allows for extrapolation to simulate lines belonging to the excited vibrational and torsional states at lower spectroscopic temperatures. For these cases, it would be extremely useful to either have experimental data at lower temperature or a QM assignment of these vibrational states.

Figure 11 shows the comparison of the experimental data (green trace), predicted ground state spectra (red dashed trace) and spectra calculated using Equation (9) and corresponding $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients. The strongest line in Figure 11 is ∼10% of the strongest line intensity in our spectral range. Since we normalize the absorbance spectrum by the experimental $nL/Q$ parameter, 77 K spectral intensities appear to be smaller than the room temperature spectral lines for the absorption cross-section spectra shown in Figure 11. The experimental noise level in the high temperature traces is amplified due to a decrease in the $nL/Q$ parameter (see Figure 6).

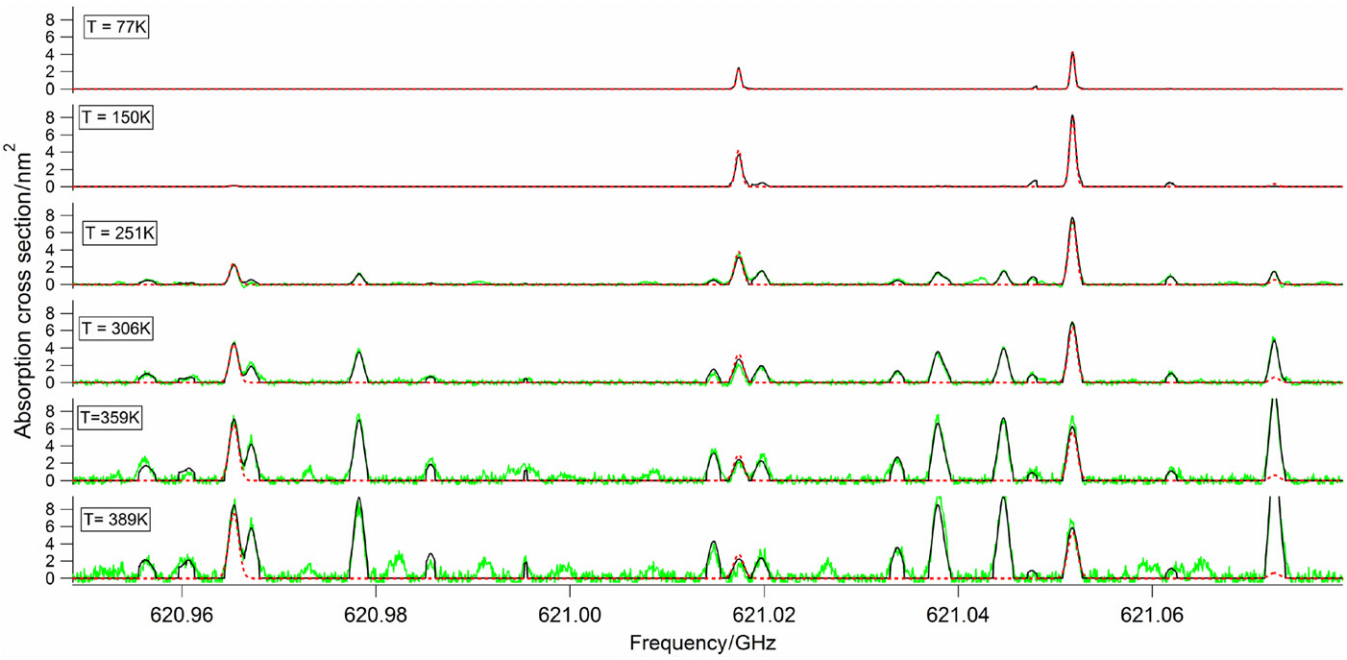


**Figure 11.** Spectra calculated using Equation (6) and corresponding $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients (black trace) for a range of spectroscopic temperatures. Green and red dashed traces are, respectively, experimental data and predictions for the ground state transitions.

(A color version of this figure is available in the online journal.)

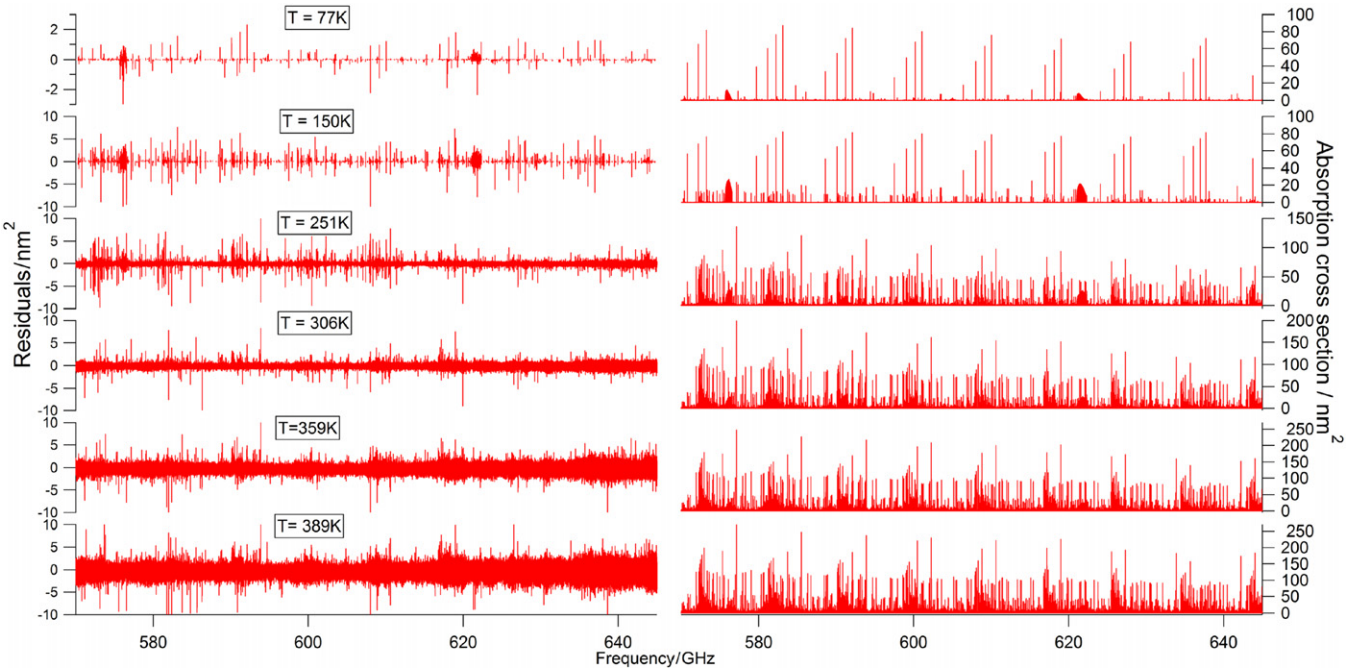


**Figure 12.** Left panel shows the residuals obtained by subtracting spectra calculated using Equation (8) and corresponding $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients from experimental (bottom four traces) and predicted ground state (top two traces) spectra for a range of spectroscopic temperatures. Right panel shows entire corresponding spectra. Residuals for 77 K and 150 K predicted spectra were only calculated for $\sim$ 1 MHz surrounding each ground state spectral line.

(A color version of this figure is available in the online journal.)

Spectral intensities as well as line shapes are reproduced very accurately for the experimental and predicted spectra (77 K and 150 K). Encouraged by the accuracy of the data reproduction shown in Figure 11, we are working on numerical algorithms that would allow us to fit multiple spectral peaks to deal with spectral blends.

To evaluate the accuracy of our fitting and prediction method, we calculated the differences between experimental (predicted) and calculated (QM) spectra for a range of spectroscopic temperatures. This is shown in Figure 12. The residuals shown in the left panel of Figure 12 are mostly within 1% of the maximum intensity that can be seen in the corresponding right panel. If we chose the rms of the residuals as the figure of merit, this would indicate an accuracy better than 0.5% in reproduction of spectral intensities, probably an optimistic result. At 77 K and 150 K where no experimental data were available, the

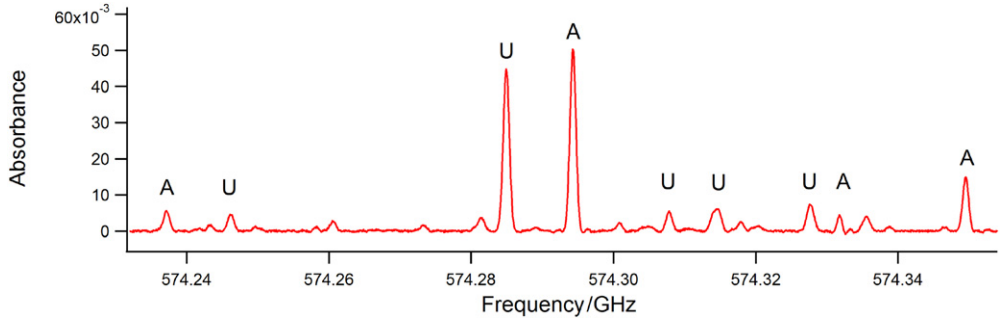


**Figure 13.** Small portion of the spectrum of ethyl cyanide at 300 K. The lines in this figure marked with an "A" are assigned in the catalog, whereas those marked with a "U" are not, but both "U" and "A" lines are among the 3355 lines in Table 1.

(A color version of this figure is available in the online journal.)

**Table 2**
Astrophysical Catalog in Strength—Lower State Energy Format

| Line | Frequency | $S_{ij}\mu^2$ | Energy | Figures | Quantum Numbers | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Type | /MHz | /D$^2$ | /cm$^{-1}$ | of Merit | J″ | $K_{-1}$″ | $K_{+1}$″ | J′ | $K_{-1}$′ | $K_{+1}$′ |
| Cat | 574237.054 | 50 | 661 | | 66 | 4 | 62 | 65 | 5 | 61 |
| Exp | 574237.212 | 42 | 624 | | | | | | | |
| Exp | 574246.215 | 186 | 973 | | | | | | | |
| Exp | 574284.959 | 1027 | 859 | | | | | | | |
| Cat | 574294.181 | 41 | 566 | | 61 | 5 | 57 | 60 | 4 | 56 |
| Exp | 574294.288 | 843 | 803 | | | | | | | |
| Exp | 574307.808 | 271 | 1025 | | | | | | | |
| Exp | 574314.427 | 2001 | 1390 | W | | | | | | |
| Exp | 574327.693 | 1825 | 1358 | | | | | | | |
| Cat | 574331.722 | 121 | 791 | | 67 | 13 | 54 | 67 | 12 | 55 |
| Exp | 574331.77 | 70 | 774 | W | | | | | | |
| Cat | 574349.414 | 1560 | 1160 | | 64 | 27 | 37 | 63 | 27 | 36 |
| Exp | 574349.513 | 1361 | 1150 | | | | | | | |

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

residuals contain a large number of narrow (∼line width) features that in part can be explained by the differences in QM and experimental line positions that become more prominent at lower temperatures where Doppler-limited line width is smaller. At 251 K, the lowest experimental temperature shown in Figure 12, we observe similar trends in the residuals which also can be explained by the inclusion of the predicted 77 K rotational spectrum, weighted 10 times higher than the rest of the data, into the fit for $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients.

## 6. RESULTS OF CATALOG FITS FOR $S_{ij}\mu^2$ AND $E$

Figure 13 shows a small portion of the spectrum at 300 K. The lines in this figure marked with an "A" belong to the ground state and are assigned in the catalog, whereas those marked with a "U" are not, but both "U" and "A" lines are among the 3355 lines analyzed in this paper. These lines are relatively weak, ranging from about 1% to 3% of the strongest lines in the spectrum for the seven weakest marked lines to 15%–20% for the two strongest lines. The threshold between the 3355 analyzed lines and the remainder of the lines in Figure 1 can be inferred by the strength of the remaining lines in Figure 13.

Table 2 shows our results in a catalog format. Because of the aforementioned spectral overlaps, all of the experimental lines and all of those calculated from a QM catalog are listed separately. In the electronic files only our experimental data are included. In the example shown here, there is a clear association between each of the four catalog lines and an adjacent line with a similar frequency.

First, consider the assigned line near 574294 MHz. While the frequency of it agrees to within 0.13 MHz of the catalog value, the observed intensity is about 20 times that of the calculated intensity, and the lower state energy difference is 237 cm$^{-1}$. This is a clear case of a circumstance in which a stronger unassigned line overlaps a weaker catalog line. Differences in lower state energy and transition strength for the other two lines are within the distributions shown in Figures 7 and 8 for lines of absorbance 0.01.

Table 3 contains index, spectroscopic temperature, and $nL/Q$ parameter for each experimental spectrum (see Figure 6).

Table 4 contains $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients (see Section 5.2) is also archived. In it $\tilde{S}_{ij}\mu^2$ is reported in units of D$^2$ (Debye squared) and $\tilde{E}$ in units of cm$^{-1}$. This file consists of 1.2 million rows, starting at 570000 MHz and incrementing in steps of 0.065 MHz. To provide for an easy way to calculate the absorbance scaled by $nL/Q$ parameter at an arbitrary temperature using archived $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients, we reformulate Equation (9) as

$$\frac{A(\nu)}{nL/Q} = C_1 \cdot \sqrt{\mathrm{M}}\frac{\left(1 - e^{-C_2 \cdot \frac{\nu}{T}}\right)}{\sqrt{T}}\tilde{S}_{ij}\mu^2 e^{-C_3 \cdot \frac{\tilde{E}}{T}}, \qquad (11)$$

**Table 3**
Spectroscopic Temperature and $nL/Q$ Parameter for Each of the 271 Scans, with the 0 Index Representing the Theoretical 77 K Spectrum Used in the Analyses

| Index | $nL/Q/(\mathrm{nm}^{-2})$ | $T/(\mathrm{K})$ |
|---|---|---|
| 0 | 1.000E+00 | 77.00 |
| 1 | 2.716E–03 | 247.62 |
| 2 | 2.707E–03 | 247.96 |
| 3 | 2.707E–03 | 247.96 |
| 4 | 2.706E–03 | 247.85 |
| 5 | 2.697E–03 | 248.10 |
| 6 | 2.687E–03 | 248.31 |

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

**Table 4**
Spectroscopically Determined $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ Coefficients

| $\tilde{S}/(\mathrm{D}^2)$ | $\tilde{E}/(\mathrm{cm}^{-1})$ |
|---|---|
| 126.3375 | 1336.774 |
| 197.5496 | 1396.586 |
| 240.094 | 1414.767 |
| 256.7859 | 1399.148 |
| 335.2192 | 1435.086 |
| 325.1441 | 1405.411 |
| 439.6971 | 1455.531 |

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

with

$$C_1 = 54.5953 \frac{\mathrm{nm^2 K^{1/2}}}{\mathrm{amu^{1/2} D^2}}, \quad C_2 = 4.799237 \times 10^{-5} \frac{\mathrm{K}}{\mathrm{MHz}},$$
$$C_3 = 1.43877506 \frac{\mathrm{K}}{\mathrm{cm}^{-1}},$$

and $M$ being the molecular weight in amu (55 amu for ethyl cyanide), $T$ temperature in K, and $\nu$ frequency in MHz.

Because this approach is based on frequency points, not spectral lines, there is no explicit line count. However, there are more than 11,000 features at 300 K with S/N greater than 2, whose absorbances are greater than 0.0007. This number is larger than the 3355 lines included in the line-by-line catalog in part because the fitting procedure effectively averages over the 271 spectra. As a result, some of the lines in this analysis are due to the $^{13}$C isotopic species.

## 7. FIGURES OF MERIT FOR CATALOG FITS

In dense spectra, both in the laboratory and in the field, issues such as blends and overlaps must be dealt with. Accordingly, we have developed a number of figures of merit that are appropriate for the catalog data of Table 2. Many of these tests are made possible because we have a very large amount of data over a range of well-defined temperatures.

*"W"—line widths*. For a given line, if the line widths calculated by the line fit for more than 50 of the 271 temperatures were more than 0.3 MHz larger than the calculated Doppler width, the line was labeled with a "*W*". Excess width is the simplest measure of a blend.

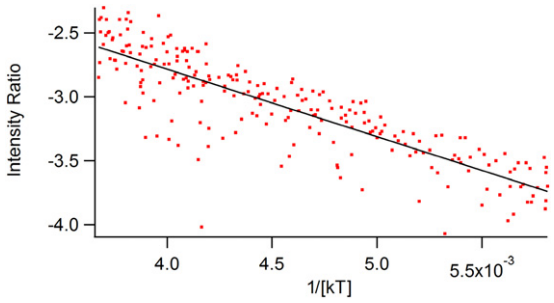


**Figure 14.** Logarithmic $1/T$ plot for a partially resolved blend.
(A color version of this figure is available in the online journal.)

*"G"—Gaussian fits*. For a given line and temperature, if the difference in the frequency obtained by the peak finder that was used to initially locate lines and the frequency returned by the subsequent fit differed by more than 0.3 MHz or if the amplitude returned by the fit was greater than 1 or less than 0.001, the fit was deemed unsuccessful and was excluded from the $1/T$ fit. For a given line, if more than 50 of the 271 temperatures were removed, the line was labeled with a "*G*". This is an indication of weak or blended lines that did not fit to a Gaussian lineshape.

*"T"—quality of the 1/T fit*. If the rms of the residual of this fit was greater than 0.1, the line was labeled with a "T". This tag is triggered by blends for which the components have significantly different lower state energies, even for the cases for which their transition frequencies are almost identical. Figure 10 shows an example of this, with the characteristic double decay in the $1/T$ plot. If the lines that contribute to this graph are close in frequency, such a feature will ordinarily pass the line width "W" and Gaussian fit "G" figures of merit. Figure 14 shows a $1/T$ plot for a partially resolved blend. Here there is a substantial scatter that results from the changing relative intensities as a function of temperature. In this case, this feature will ordinarily fail the line width and frequency figures of merit as well.

*"X"—exclusion test*. For a given line, if more than 200 of the 271 temperatures were excluded from the $1/T$ fit, the line was labeled with an "X". This is an indication that the line is extremely weak or severely blended at most temperatures. It is also an indication that the $1/T$ fits are based on few points, and the resulting strength and lower state energy measurements are highly suspect.

## 8. DATA ARCHIVING

Much of the success of QM models has been based on the fact that they are extensible and that, in general, the experimental data on which they are based are available in the archival literature. While there has been some improvement in frequency measurement accuracy with modern synthesis and computational techniques, for many decades SMM/THz lines have been measured to a small fraction of the Doppler limit. This archival availability was facilitated because the data consisted of a relatively small number of highly accurate measured frequencies, which made their publication straightforward, even before the advent of large electronic supplements to the journals.

**Table 5**
Experimental Absorbance Spectrum of Ethyl Cyanide

| Absorbance |
|---|
| 1.289081e–03 |
| 1.736081e–03 |
| 1.987339e–03 |
| 2.285944e–03 |
| 2.242617e–03 |
| 2.246547e–03 |
| 2.395048e–03 |

**Note.** The complete set of data is available as a tar.gz supplementary data file in the electronic edition of the journal.

In recent years, the amount of data produced by SMM/THz spectrometers has increased dramatically, but with the advent of electronic supplements to archival journals, the traceability to the original data can be maintained.

### *8.1. Data and Analysis for the Quantum Model*

Our results depend directly upon the results of QM models. For the experimental results reported here to be completely traceable and extensible, they must be based on QM models which also have these properties and which are ideally easily available. Accordingly, we have used the publicly available JPL and Cologne catalogs (for ethyl cyanide, they are functionally equivalent), which are very closely related to the published papers (Brauer et al. 2009; Fukuyama et al. 1999; Pearson et al. 1994). While the exact details of the calculations that are the bases of the catalogs are often not known, the possible variation will have little effect on the results presented here because the QM calculations are only used to establish the temperature and concentration of the experimental sample. However, for the validation of the method presented here, the quality of the available QM model was very important and is a major reason that we chose ethyl cyanide for this initial work.

### *8.2. The Experimental Data*

Just as for QM models, the experimental models are dynamic entities that should be updatable in the future. To cite only a few circumstances: (1) when QM models for low lying vibrational states become available, this can improve the accuracy of the intensities of these states at low temperature; (2) collisional cooling can be used to directly extend the range of temperature measurement to temperatures below 240 K; (3) the data analysis procedures that involve the peak finding and analysis of blended lines could be more sophisticated; and (4) the work can be extended to weaker lines.

Accordingly, we have included the data from each of the 271 variable temperature data runs as a supplementary data file in the electronic edition of the journal. Because the removal of the baseline and the power normalization are rather technical and because the resulting intensity-calibrated spectra as in Figure 1 provide a means for the extensions discussed in the previous paragraph, these 271 scans are stored in their intensity-calibrated form. Each text file consists of 1.2 million absorbances, starting at 570000.0 MHz and incrementing in steps of 0.0625 MHz (a small portion of one of these files is shown in Table 5). Each file name contains the experimental index of the spectrum (in the order the scans were recorded) together with spectroscopically determined temperature. To maintain the traceability of our analysis, we also provide the predicted 77 K spectrum used in fitting for $\tilde{S}_{ij}\mu^2$ and $\tilde{E}$ coefficients (see Section 5.2). It is archived with index 0 and $nL/Q$ parameter set to 1.

## 9. SUMMARY

This paper demonstrates that the experimental observation of intensity-calibrated SMM/THz spectra at many temperatures can dramatically increase the completeness of the astrophysical database. More importantly, the resulting catalog is complete in the sense that the strongest lines are included as opposed to simply the lines that are easiest to calculate. In addition, since the results of this work include the lower state energy and transition strength associated with each measured transition frequency, we expect that these data will be extremely useful in the QM assignment and analysis of as yet unanalyzed states and transitions.

An implementation is shown that provides a means to connect the basis for this experimental approach to ongoing QM analyses that is extensible as spectroscopic knowledge grows and expands and that provides a convenient interface for the astrophysical community. It is likely with the growing power and complexity of SMM/THz telescopes and a similar expansion of the underlying spectroscopy that additional instrument specific tools will become advantageous.

We would like to thank the National Science Foundation and JPL/Herschel for their support of this work. This work was also supported by NASA Headquarter under the NASA Earth and Space Science Fellowship Program—Grant NNX09AP10H.

## REFERENCES

Brauer, C. S., Pearson, J. C., Drouin, B. J., & Yu, S. 2009, ApJS, 184, 133
Fukuyama, Y., Omori, K., Odashima, H., Takagi, K., & Tsunekawa, S. 1999, J. Mol. Spectrosc., 193, 72
Gordy, W., & Cook, R. L. 1984, Microwave Molecular Spectra, Techniques of Chemistry Vol. 18 (3rd ed.; New York: Wiley)
Heise, H. M., Lutz, H., & Dreizler, H. 1974, Z. Naturf a., 29, 1345
Laurie, V. W. 1959, J. Chem. Phys., 31, 1500
Medvedev, I., Winnewisser, M., De Lucia, F. C., Herbst, E., Bialkowska-Jaworska, E., Pszczolkowski, L., & Kisiel, Z. 2004, J. Mol. Spectrosc., 228, 314
Medvedev, I. R., & De Lucia, F. C. 2007, ApJ, 656, 621
Mehringer, D. M., Pearson, J. C., Keene, J., & Phillips, T. G. 2004, ApJ, 608, 306
Müller, H. S. P., Schlöder, F., Stutzki, J., & Winnewisser, G. 2005, J. Mol. Struct., 742, 215
Pearson, J. C., Sastry, K. V. L. N., Herbst, E., & De Lucia, F. C. 1994, ApJS, 93, 589
Petkie, D. T., Goyette, T. M., Bettens, R. P. A., Belov, S., Albert, S., Helminger, P., & De Lucia, F. C. 1997, Rev. Sci. Instrum., 68, 1675
Pickett, H. M., Poynter, R. L., Cohen, E. A., Delitsky, M. L., Pearson, J. C., & Muller, H. S. P. 1998, J. Quant. Spectrosc. Radiat. Transfer, 60, 883
Remijan, A. J. 2009, http://www.splatalogue.net/
Townes, C. H., & Schawlow, A. L. 1955, Microwave Spectroscopy (New York: McGraw-Hill)